All-Electron Wavefunction of Electron Transfer Protein Cytochrome *c* by Density Functional Theory

Fumitoshi Sato, Tamotsu Yoshihiro, Makoto Era & Hiroshi Kashiwagi
Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, Iizuka, Fukuoka 820-8502, Japan

Cytochrome *c* (cyt. *c*), which exists in the respiratory chain in the inner mitochondrial membrane, is a famous electron transfer protein. The horse heart cyt. *c* contains 104 residues and a *c*-type heme (Fe-protoporphyrin). Protoporphyrin is bound covalently to polypeptide (apo-protein) of cyt. *c* via thioether bonds between two vinyl side chains and Cys 14 and 17. Residues His 18 and Met 80 form the 5-th and 6-th ligands of heme iron, respectively.

One of the direct methods to understand the reactivity of cyt. *c* is to obtain the electron wavefunction of cyt. *c* quantitatively. Quantum chemical calculations on proteins are coming within range with recent molecular orbital (MO) methods as very large organic molecules [1-4]. Some of them are based on the localized orbital method that does not provide general solutions, and the others depend on the *ab initio* Hartree-Fock (HF) equation that does not include the electron correlation effect. However, calculations on metalloproteins require precise methods containing the suitable treatment for electron correlation such as density functional theory (DFT).

In this study, in order to reveal the characteristics of the electronic state in a metalloprotein, we carried out a computation of an all-electron wavefunction of horse heart d^6-low-spin ferrocytochrome *c* (ferrocyt. *c*) [5] by our Gaussian-based DFT MO program, ProteinDF [6]. It may be the first full-scale DFT calculation of a metalloprotein. The numbers of atoms, electrons, orbitals and auxiliary functions are 1,738, 6,586, 9,600 and 17,578, respectively.

MOs of d^6-low-spin ferrocyt. *c* are delocalized over the whole molecule, and it is no exception for the orbitals which 3*d* orbitals of Fe participate in. We show the 3 dimensional (3D) graphics of the 3293-th highest occupied MO (HOMO) with cyt. *c* structure in Figures 1A-D, where the isosurface values are (A) ±0.05, (B) ±0.005, (C) ±0.0005 and (D) ±0.00005, respectively. The scale of Fig. 1A is magnified by twice. As can be seen in Fig. 1A, the main components of HOMO are 3*d* orbitals of Fe, and the gross atomic orbital population indicates that the components of d_{xy}, d_{xz}, d_{yz} and the total of *d* orbitals occupy 65.4, 10.5, 9.3 and 85.2 %, respectively. Those of d_{x2-y2}, d_{3z2-r2} are under 0.1 %.

This MO is considered as the carrier bag in electron and hole transfer, since the task of transferring electron is undertaken by oxidation/reduction reactions of heme iron. It is delocalized over the whole molecule and attains the outside of cyt. *c* with almost keeping the phase of d_{xy} within dozens of Å (Fig. 1A-C). HOMO spreads out from Fe and is relatively off center within the molecular structure (Figs. 1C and D). It is obvious that the extent of HOMO is abnormally broad to compare with 3*d* Slater type atomic orbitals of Fe. At long distance electron transfer in metalloproteins, it has been considered that the direct coupling between cyt. *c* HOMO and acceptor orbitals is extremely small. However, Fig. 1 suggests that the direct coupling between cyt. *c* HOMO and acceptor MOs cannot be ignored in the processes of electron transfer in cyt. *c*.

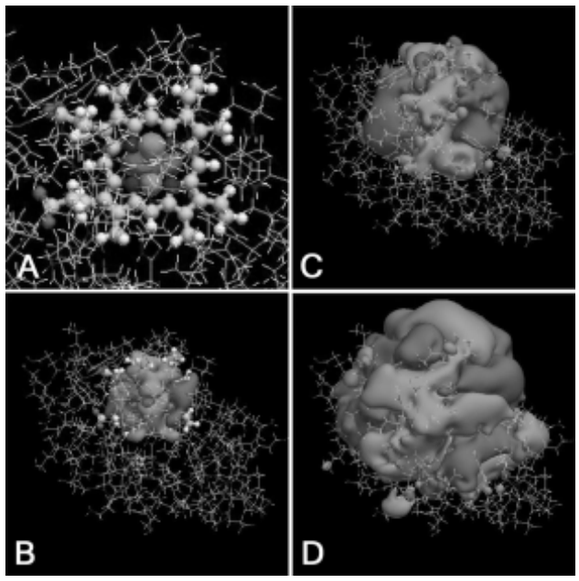

Figure 1: The 3D graphics of the 3293-th MO (HOMO) in d^6-low-spin ferrocyt. *c*. The isosurface values are (A) ±0.05, (B) ±0.005, (C) ±0.0005 and (D) ±0.00005, respectively. Dark and bright isosurfaces indicate plus and minus regions of MO, respectively. The scale of (A) is magnified by twice.

Our computational data of this study give important information to the further MO calculations for larger proteins. The calculation was carried out with workstation cluster composed of 15 Alpha workstations and converged with 59 SCF iterations. The total elapse time for the first iteration of SCF calculation was 74,676 sec (20.74 hrs) in cyt. *c*. The slope between the elapse time of total calculation and the number of orbitals in the parts of and whole of cyt. *c* by logarithm-logarithm plot is 2.4, which gives us one of the standards to compute the larger proteins. Nowadays some computer centers contain about 1,000 times larger resources than those used in this study, so that it is possible to calculate the 10 times larger molecules (100,000 orbitals). Then, important membranous proteins such as photosynthetic proteins are in the range of computations.

Acknowledgement
This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture and from the Science and Technology Agency of Japan.

References

[1] J. J. P. Stewart, *Int. J. Quant. Chem*. **58** (1996) 133.
[2] M. Challacombe, E. Schwegler, *J. Chem. Phys*. **106** (1997) 5526.
[3] P. M. W. Gill, *Chem. Inf. Comp Sci. Chem. Soc. Japan* **17** (1999) 4.
[4] K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, *Chem. Phys. Lett*. **312** (1999) 319.
[5] F. Sato, T. Yoshihiro, M. Era, H. Kashiwagi, *Chem. Phys. Lett*. (2001) in press.
[6] F. Sato, Y. Shigemitsu, I. Okazaki, S. Yahiro, M. Fukue, S. Kozuru, H. Kashiwagi, *Int. J. Quant. Chem*. **63** (1997) 245.